Empirical Study of Design-Construction Feedbacks in Building Construction Projects

Kiavash Parvan[a], Hazhir Rahmandad[b], Ali Haghani[c]

(a) PhD in project management program, Civil and Environmental Engineering Department, University of Maryland, kparvan@umd.edu
(b) Associate professor, Industrial and Systems Engineering Department, Virginia Polytechnic Institute and State University, hazhir@vt.edu
(c) Chair and professor, Civil and Environmental Engineering Department, University of Maryland, Haghani@umd.edu

Abstract

Understanding project dynamics is one of the core application areas for system dynamics. Despite a long tradition of modeling the interactions between multiple phases in a project model, the strength of these feedback mechanisms have not been rigorously estimated. In this article we take a step towards addressing this shortcoming by estimating the feedbacks between design and construction phases of construction projects. We estimate the parameters of three hypothetical feedback relations between design and construction with data from 15 construction projects. Consistent with previous qualitative evidence, the estimated factors reveal that undiscovered design rework diminishes construction quality and production rate significantly and construction completion speeds up the detection of undiscovered design rework. We also assess the predictive power of our model using another set of 15 empirical cases. The model showed excellent fit to the calibration data calibration and good prediction in validation.

Problem statement

Project modeling has a long history in the system dynamics literature. Starting with a model that informed the arbitration of a ship building project lawsuit (Cooper 1980), this line of modeling has grown to one of the most successful areas of system dynamics practice (Lyneis and Ford 2007). While the rework cycle is at the core of project models, from early on the modelers identified the importance of disaggregating these models to include multiple phases or task groupings (Lyneis and Ford 2007). Formulating the multi-phase project models were discussed in detail by (Ford & Sterman, 1998) in the semiconductor industry and many applications have used different variants of this formulation ((Khoueiry, Srour, & Yassine, 2013; Lee, Han, & Peña-Mora, 2009; Park, Kim, Lee, & Han, 2011)). In this formulation each phase of the project is modeled separately, with the knock-on effects of the quality and progress of each phase on the successive phases. Different effects could be conceived in this set up, the most prominent of which are the impact of early phase quality on later phase productivity, the effect of early quality on later quality, and the effect of later completion of tasks on the discovery of errors in earlier phases. These effects could then activate endogenous rework, schedule pressure, and morale loops within different phases, leading to much variability in project performance, quality, and costs (Ford and Sterman 1998; Lyneis and Ford 2007). However, the strength of these feedback mechanisms has been assumed based on qualitative knowledge of each case, and rigorous empirical estimates are lacking in the literature. This shortcoming in the literature may be partially due to the complexities of collecting time series data required for such estimation tasks and partially due to the oneoff nature of many dynamic models of projects which limits the statistical power which can be expected from estimation.

The latent impact of design error on construction phase has been studied by some statistical research. Baruti and colleagues (Burati, Farrington, & Ledbetter, 1992) reported that design failure or defect is responsible for 79% of the total change costs, and 9.5% of the total project cost. Cusack (Cusack, 1992) showed that documentation errors increase project costs 10%. Hanna and colleagues (Hanna, Camlic, Peterson, & Nordheim, 2002) found that design errors lead to 38%-50% of change orders in the projects under their studies. An recently, Lopez and Love (Lopez & Love, 2012) showed that the average of direct and indirect cost for design errors is about 7% of contract value. Nevertheless such estimates at a level of aggregation useful for dynamic modelers are lacking. In light of the important roles these feedback effects play in typical project models, a more reliable quantitative estimate will strengthen practical models for project planning and project dispute resolution, and provide better grounding for future theoretical work.

Methods and Data Overview

In this study, we quantify the design-construction feedback relationships in design-bid-build (DBB) construction projects[1]. A generic dynamic model with two phases of design and construction is developed based on the SD literature. Historical data from 30 building construction projects is used to estimate and validate the model. The model is calibrated with 15 randomly selected projects and the other 15 projects are used

[1] DBB is a project delivery method which design and construction are performed in two separate phases with no overlap. Design-Build (DB) and Construction Management (CM) are the other examples.

for validation. The calibration process is used to estimate three distinct effects: 1) impact of design quality on construction quality, 2) the effect of design quality on construction productivity, and 3) the effect of construction progress on error discovery rate in design. The validation process informs the feasibility of using simple SD models to estimate the likely distribution of project outcomes for new projects, a key step in project planning activities.

The dataset includes, for each project, the (initially) estimated duration (duration based on planning), estimated cost, actual duration, actual cost, and the cost trajectory of project over time based on owner payments, all separable by the design and construction phases. The sample statistics for estimated time to finish (F_0), the ratio of actual to estimated time to finish (F/F_0), estimated cost (W_0) and the ratio of actual to estimated cost (W/W_0) are shown in Table 1 for design (D) and construction (C) phases of calibration and validation projects.

Table 1: Descriptive statistics of calibration (n=15) and validation (n=15) data

Sample	Variable	Unit	Mean	Std Dev	Median	Minimum	Maximum
Calibration	D_F0	Month	10.9	7.70	5.83	3.5	26.3
	D_F/F0	Dmnl	1.4	0.37	1.22	1.0	2.1
	D_W0	$	1.48E+06	1.59E+06	6.00E+05	6.90E+04	4.42E+06
	D_W/W0	Dmnl	1.3	0.31	1.21	0.9	2.2
	C_F0	Month	32.0	18.39	31.10	9.4	65.0
	C_F/F0	Dmnl	1.4	0.47	1.42	0.9	2.6
	C_W0	$	1.81E+07	2.06E+07	6.35E+06	3.48E+05	5.51E+07
	C_W/W0	Dmnl	1.1	0.29	1.05	0.7	1.8
Validation	D_F0	Month	12.0	10.81	10.43	3.3	47.2
	D_F/F0	Dmnl	1.2	0.26	1.07	1.0	1.7
	D_W0	$	1.12E+06	1.27E+06	5.75E+05	1.12E+05	4.06E+06
	D_W/W0	Dmnl	1.1	0.19	1.11	0.9	1.7
	C_F0	Month	27.7	16.64	24.97	11.0	68.6
	C_F/F0	Dmnl	1.5	0.39	1.43	1.0	2.1
	C_W0	$	1.21E+07	1.43E+07	6.37E+06	3.48E+05	4.23E+07
	C_W/W0	Dmnl	1.1	0.07	1.07	0.9	1.2

Model development

The construction project model is developed at the level of design and construction phases. In each phase, the completion of tasks was followed by a review process (called “design review” and “inspection” in the design and construction phases respectively (Figure 1)). The conjunction of the work and review activities is modeled utilizing the simple rework cycle concept developed by (Richardson G. P. and Pugh, 1981). While more complex rework cycle formulations exist (e.g. see (Ford & Sterman, 1998; Rahmandad & Hu, 2010)) the simple 3-stock formulation is consistent with the level of aggregation available from our data, which does not include details on individual tasks or rework items, and therefore appropriate for the current application.

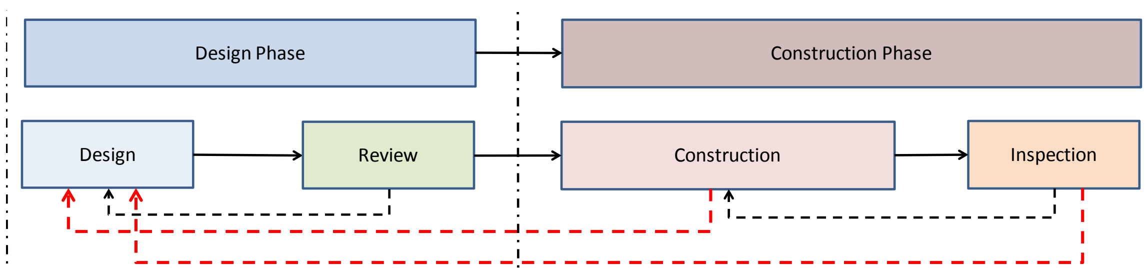

Figure 1: Construction project work flow

Figure 2 overviews the model developed in Vensim. The model captures two phases of design and construction in two separate rework loops. The rework loop parameters of production rate (P), error rate (E) and time to detect undiscovered rework (D), are normalized by project initial values i.e. initial work (W0) and Duration (T0) to be comparable across different projects.

In reality the start and finishing of each DBB project are regulated by five events: 1) Design Start, 2) Construction Document (CD) Finish, 3) Construction Start, 4)

Construction Finish, and 5) Design Service (DS) Finish. Design Start, "*D Start*", is the event that initiates design. Design finishes when construction documents (CD) are approved and delivered for bidding process. We track the time at which the design phase is perceived to be complete by the variable "*D CD Finish*". We assume the design phase is completed when the approved design work passes a threshold (99%) of initial design work. Design CD finish triggers the start of the bidding process, during which neither design nor construction activities progress. The next event, the construction start ("*C_Start*"), commences at the end of the bidding process. Construction proceeds until "*Construction finish*" event occurs. "*Construction finish*" event is triggered where the construction "*Approved work*" passes a threshold (99%) of initial construction work. Meanwhile, some design reworks/errors may remain undiscovered by the design finish. These are eventually discovered and fixed during construction phase. In DBB projects, usually the same architectural and engineering (A/E) designer is enrolled to provide design services (DS) during construction phase, therefore the initial design and later design services could be seen as the same process and are represented by a single stock and flow diagram. The last event is "*D DS Finish*". We assume the design services during construction, "*D DS Finish*", is completed when the approved design work passes a threshold (99%) of initial design work and undiscovered design rework stock level is less than (1%). We use our data to specify the "*Design Start*" and "*Construction Start*" events for each project, while the "*D CD Finish*", "*Construction finish*", and "*D DS Finish*" are all endogenously calculated.

Phase inter-relationship between design and construction has been studied by many. Several hypotheses have been proposed to describe the design-construction interaction mechanisms. Some researchers have proposed the design rework/error as the main contributor to (lack of) construction quality. (Lyneis & Ford, 2007) call this as "errors build errors" effect which "*undiscovered errors in upstream work products (e.g., design packages) that are inherited by downstream project phases (e.g., construction) reduce the quality of downstream work as these undiscovered problems are built into downstream work products*". They also cite the works of Pugh–Roberts Associates (PRA), (Abdel-Hamid TK, 1984; Ford et al., 2004; Lyneis, Cooper, & Els, 2001), Ford-Sterman, and Strathclyde models as examples. Some others have proposed design change as the main contributor to reduce construction labor productivity (Hanna, Asce, & Gunduz, 2004; Hanna et al., 2002; Hanna, Russell, Gotzion, & Nordheim, 1999; B. C. W. Ibbs, 1997; W. Ibbs, 2005; Moselhi, Assem, & El-Rayes, 2005). Building on these theoretical motivations we conducted five expert interviews with three senior project managers, in different positions to better understand factors influencing quality and productivity in design and construction phases. Three mechanisms were identified:

1. Undiscovered design rework may increase construction error rate
2. Undiscovered design rework may slow down construction production rate
3. Construction progress may increase the detection rate of undiscovered design reworks

These mechanisms are consistent with the previous system dynamics research (Lyneis & Ford, 2007). We therefore model three feedback mechanisms between design and construction phases. We capture the first knock-on mechanism, "*Factor>A*", in Equation 1. We assume the construction error rate is a function of the undiscovered rework in the design phase multiplied by a project-specific parameter representing base construction error, Equation 2. The design undiscovered rework, "*D UndiscoveredRework*", is divided by design initial work, "*D W0*"[2], for being scaled between -1 and +1. We allow for negative rework to capture scope reduction in construction.

$$\text{"Factor A"} = (1+\mathrm{Max}(0, \text{"D UndiscoveredRework"}/\text{"D W0"}))^{\wedge}a \qquad (1)$$

$$\text{"C InfluencedErrorRate"} = \mathrm{MIN}(1, \text{"C E"} * \text{"Factor>A"}) \qquad (2)$$

In the absence of data on human resources allocated to the project, in each phase a single productivity parameter is used to capture both the number of project employees and the productivity per full-time equivalent (FTE). While this factor, "*D P*", is assumed constant in the design phase for each project (but different across different projects), the construction work rate is impacted by the undiscovered rework in the previous (i.e. design) phase, through the *"Factor>B"* effect (Equation 3). This is the second knock-on effect that we capture in our model. Equation 4 demonstrates how we normalize model parameter production rate (P), to be estimated in calibration, by initial work (W0) and estimated work duration (T0).

$$\text{"Factor B"} = (1-\mathrm{Max}(0, \text{"D UndiscoveredRework"}/\text{"D W0"}))^{\wedge}b \qquad (3)$$

[2] Variable names start with the phase (C for Construction and D for Design) followed by the descriptive concept.

"C Work rate" = "C W0" / "C T0" * "C P" * "Factor>B" (4)

Rework discovery is assumed to happen through a first order draining from the stock of undiscovered rework. The time constant for this delay is set as another project-specific constant (D) for the construction phase. However we assume the construction progress allows faster discovery of design problems and therefore will reduce the time constant for rework discovery in design phase, the third inter-phase factor we model (Equation 5). Equation 6 shows how the model parameter time to detect rework (D) is normalized and how Factor C influences design rework detection rate.

"Factor C" = 1/(1+"C AcceptedWork"/"C W0")^c (5)

"D Detection rate" = "D UndiscoveredRework" / ("D D" *"D T0"*"Factor C") (6)

Figure 2 provides an overview of the causal relationships in the model. The switches and variables that regulate the timing of activation of different phases are not shown for clarity. Parameters that are calibrated are highlighted in bold and larger font, with exogenous variables such as initial scope and schedule in underlined italics. Full model documentation, following minimum model documentation guidelines (Rahmandad & Sterman, 2012) is available in an online appendix with the complete simulation models.

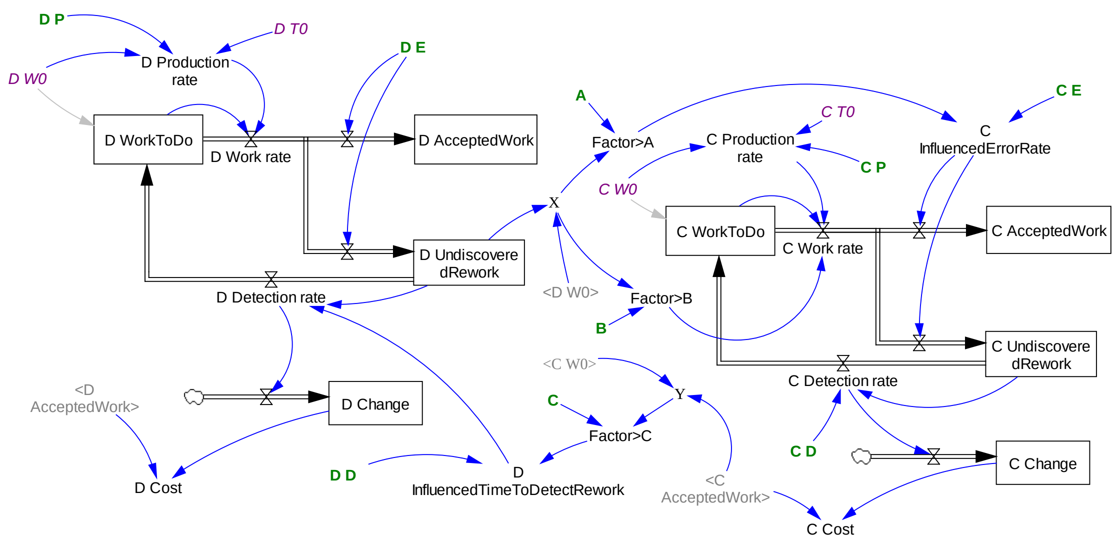

Figure 2: Overview of the model causal structure

Model calibration

Our goal in the model calibration is to estimate the parameters of our generic model to best represent the 15 randomly selected projects assigned for calibration. The model calibration result has two applications. First, we are able to find the range and distribution of project-specific parameters (i.e. error rate (E), production rate (P), and time to detect rework (D) for the two phases of design and construction). This information can then be used to form expectations on these parameter values when facing the task of estimating a new project. Second, we want to estimate the three inter-phase feedback effects (design quality on construction quality, design quality on construction productivity, and construction progress on design rework discovery rate). This information is valuable both theoretically, and for practical project planning purposes.

Calibration is typically conducted as a numerical optimization to estimate model parameters, minimizing the error between the model outputs and data (Oliva 2003). In our project we define the objective (payoff) function to be minimized as a linear combination of three error components; the squared percentage errors of time, total cost and cost curve, summed over both phases. Figure 3 illustrates the payoff function components.

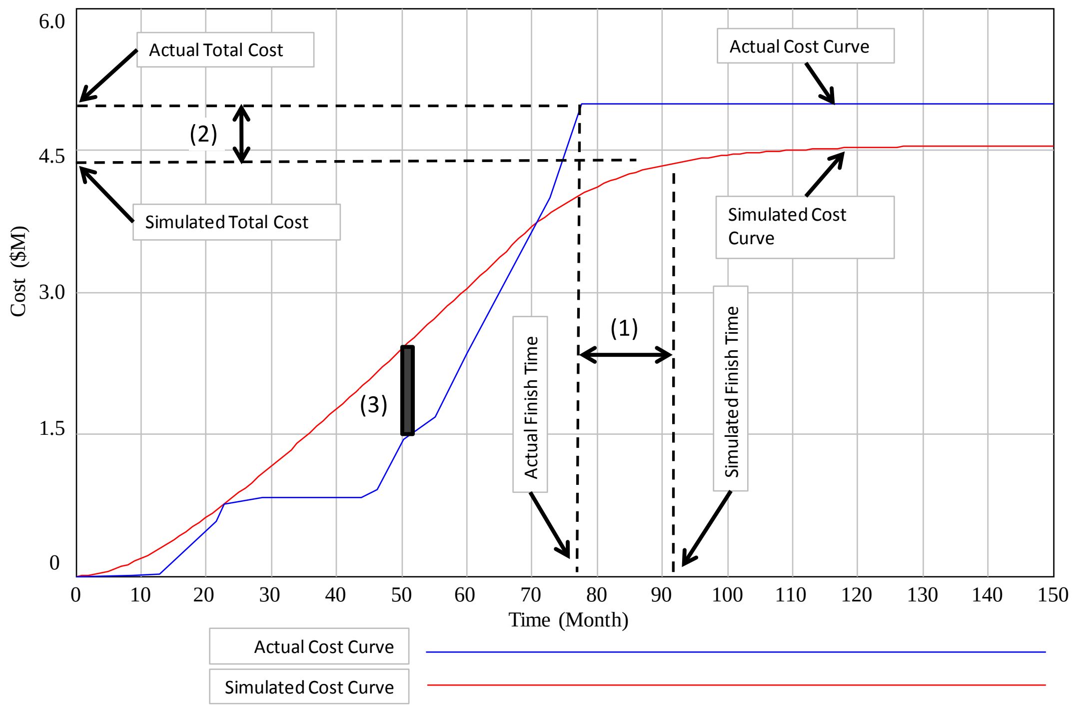

Figure 3: Calibration payoff function components

Equation 7 and Equation 8 formulate the payoff functions of design and construction respectively. Equation 7 includes four elements for the design phase: 1) the squared percentage[3] error of design construction document finish (D_CD), 2) the squared percentage error of design services during construction (D_DS), 3) the squared percentage error of design total cost (D_CT) and 4) the squared percentage error of design cost curve (D_CC(t)). Equation 8 formulates the construction payoff function in

[3] In calculating the percentages we use the average of actual and simulated in the denominator. This avoids division by zero early in the calibration process while keeping the payoff function robust. The alternative formulation that includes only the actual values in the denominator makes not qualitative difference in the results but leads to more computational errors.

the same manner, except that the construction payoff function has only one component for time, which is construction finish time (C_F).

$$Payoff(Design) = \sum_i \left\{ W_{D\_CD} \left(\frac{D\_CD_{sim,i} - D\_CD_{act,i}}{\left|D\_CD_{sim,i}\right| + \left|D\_CD_{act,i}\right|} \right)^2 + W_{D\_DS} \left(\frac{D\_DS_{sim,i} - D\_DS_{act,i}}{\left|D\_DS_{sim,i}\right| + \left|D\_DS_{act,i}\right|} \right)^2 + W_{D\_CT} \left(\frac{D\_CT_{sim,i} - D\_CT_{act,i}}{\left|D\_CT_{sim,i}\right| + \left|D\_CT_{act,i}\right|} \right)^2 + W_{D\_CC} \frac{1}{D\_Dur_{sim,i}} \int_0^{D\_Dur_{sim,i}} \left(\frac{D_{CC\,sim,i}(t) - D_{CC\,act,i}(t)}{\left|D_{CC\,sim,i}(t)\right| + \left|D_{CC\,act,i}(t)\right|} \right)^2 dt \right\} \tag{7}$$

$$Payoff(Construction) \tag{8}$$

$$= \sum_i \left\{ W_{C\_F} \left(\frac{C\_F_{sim,i} - C\_F_{act,i}}{|C\_F_{sim,i}| + |C\_F_{act,i}|} \right)^2 \right.$$

$$+ W_{C\_CT} \left(\frac{C\_CT_{sim,i} - C\_CT_{act,i}}{|C\_CT_{sim,i}| + |C\_CT_{act,i}|} \right)^2$$

$$\left. + W_{C\_CC} \frac{1}{C\_Dur_{sim,i}} \int_0^{C\_Dur_{sim,i}} \left(\frac{C_{CC\,sim,i}(t) - D_{CC\,act,i}(t)}{|D_{CC\,sim,i}(t)| + |D_{CC\,act,i}(t)|} \right)^2 dt \right\}$$

The errors are normalized into percentages so that they could be linearly combined using weights which represent the relative importance of different components. These weights are specified subjectively based on the researchers' relative confidence in the precision of the data and the amount of information they embed. For example cost curve errors are calculated based on multiple data points (based on monthly data) which conceptually captures more information about cost variation whereas final time and final cost are a single number (fewer data points). However, as the cost curve was retrieved from the project invoice log, their level of precision is less than ideal. Therefore we reduce the weight for the cost curve and increase it for the final time and cost. Consequently the following weights are used in the calibration results reported here: $W_{D\_CD} = \frac{1}{3}, W_{D\_DS} = \frac{1}{3}, W_{D\_CT} = \frac{1}{6}, W_{D\_CC} = \frac{1}{6}$, and $W_{C\_F} = \frac{1}{2},\ W_{C\_CT} = \frac{1}{4}, W_{C\_CC} = \frac{1}{4}$. Finally, the design and construction payoff functions are combined with equal weights

($W_D = \frac{1}{2}, W_C = \frac{1}{2}$), to construct the total payoff to be minimized (See Equation 9). We perform some sensitivity analysis on the assumptions regarding the weights for the payoff function and find little substantial differences in insights within reasonable ranges for these parameters (See the section "Robustness of Calibration Results").

$$Payoff = W_D\ Payoff_{Design} + W_C\ Payoff_{Construction} \tag{9}$$

For calibration, 15 projects out of the 30 projects are randomly selected. Each project is simulated separately in the model. However, to maximize the statistical power in estimating the inter-phase feedback effects, we assumed the parameters for those effects, a, b, and c, are common industry-wide and thus are the same across these 15 projects. Therefore the 15 projects are linked together through these parameters and this requires simultaneous estimation of all projects (rather than one-by-one estimation). As a result, we classify the model parameters into two categories: 1) project-specific parameters which are independent from one project to another, and 2) industry parameters which are common across all projects. The project-specific parameters consist of production rate (P), error rate (E) and time to detect undiscovered reworks (D) for each phase (a total of 6 parameters for each project), while the industry parameters include a, b and c. Calibration was conducted in Vensim DSS 5.8 by simultaneously estimating the project-specific and industry parameters over 15 calibration projects, leading to a total of 93 (=15*6+3) parameters to be estimated.

The large parameter space required us to perform the calibration in three phases. In the first phase, we conducted a global search with multiple start points in the parameter space using a course time step (TS=0.25) and relatively large fractional tolerance of 0.003. In the second phase we first fixed industry parameters a, b, and c (using values from step 1) and optimized the model, project by project, with project specific parameters P, E, and D (15 separate calibrations). Then we fixed project specific parameters P, E, and D and optimized the model on all projects with industry parameters a, b, and c. These steps were repeated iteratively until we converged. In phase three, we switched back to more precise time step of 0.0625 and fractional tolerance of 3E-5 to find tune the optimal point found approximately in phase two. For more details, please see model documentation in the online appendix.

Following this procedure, industry parameters were estimated as a=2.169, b=2.232 and c=1.104. Table 2 shows the mean vector, standard deviation vector and correlation matrix of the project specific calibrated parameters.

Table 2: Descriptive statistics and Correlation matrix of calibrated parameters

	Mean	StdDev	D_P	D_K	D_D	C_P	C_K	C_D
D_P	0.95	0.23	1.00					
D_K	0.20	0.16	0.45	1.00				
D_D	1.62	1.54	-0.19	-0.60	1.00			
C_P	0.87	0.42	0.33	0.30	-0.15	1.00		
C_K	0.02	0.21	-0.06	-0.52	0.06	-0.17	1.00	
C_D	0.38	0.49	0.09	-0.10	-0.05	-0.54	-0.03	1.00

Figure 4 and Figure 5 show the absolute percent error (APE) of time finish, final cost and cost curve of design and construction, respectively, for the 15 projects used in validation . "*D_Valid*" and "*C_Valid*" are the weighted average errors linearly combined with the same weights used in the payoff function. The sequence of projects on horizontal axis is based on these two values sorted in descending order. Figures 6 and 7 depict two (2) examples of calibrated projects.

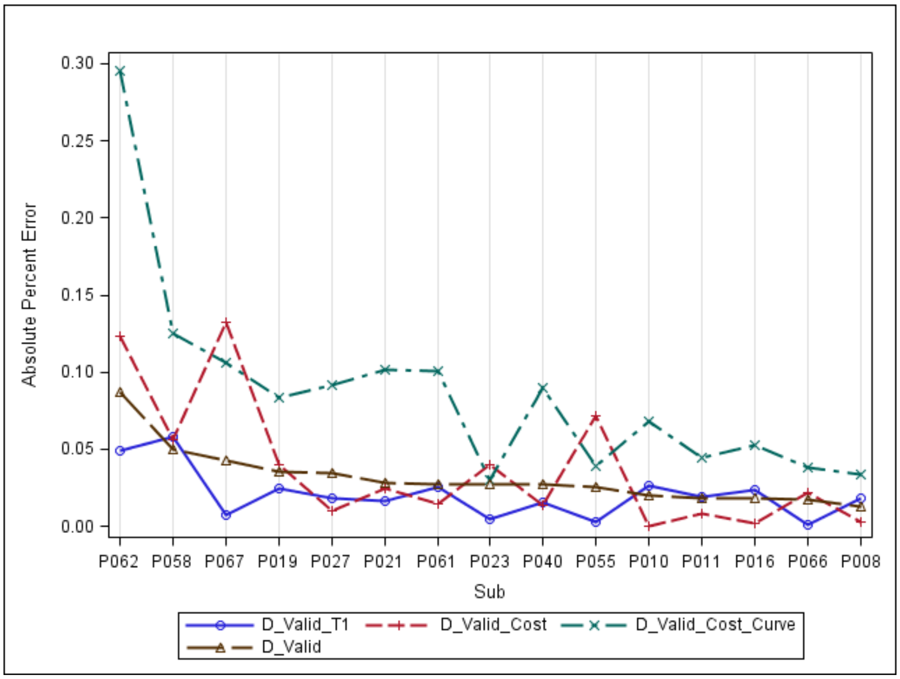

Figure 4: Design calibration error

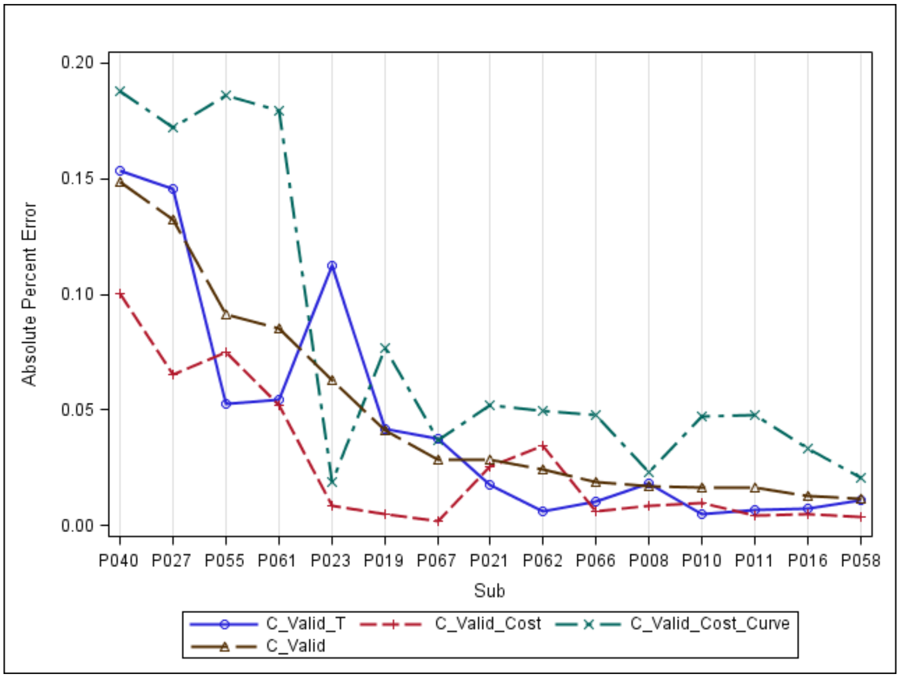

Figure 5: Construction calibration errors

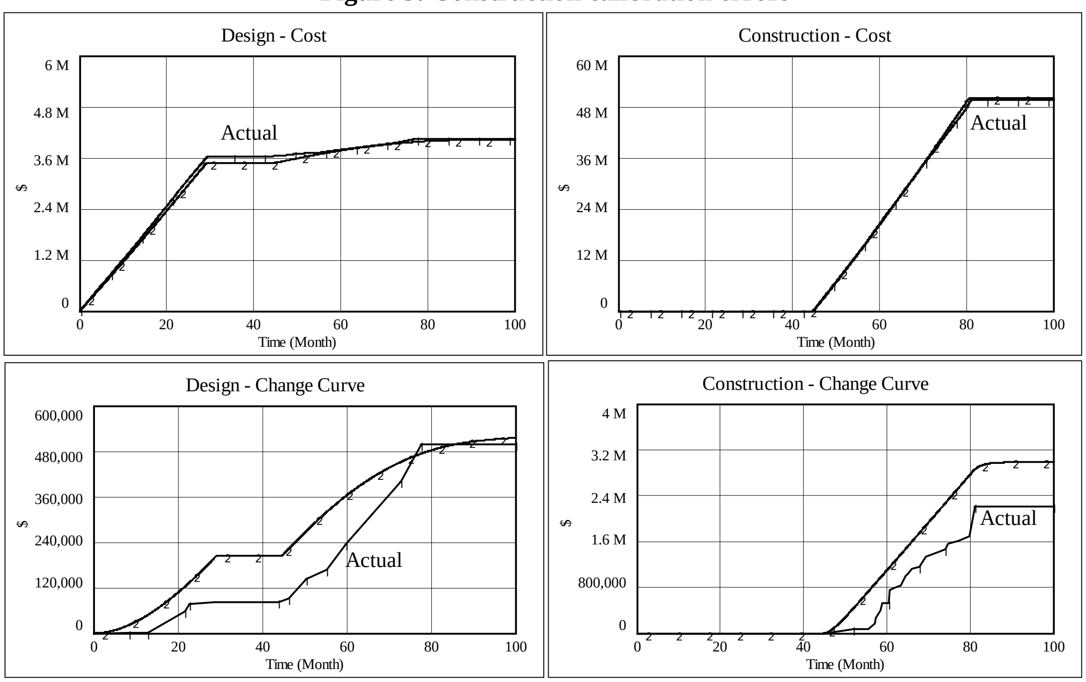

Figure 6: Simulation result of Project P008 (Best fit). Design CD Finish = 29.1 (Simulated), 29.4 (Actual). Construction Finish = 80.5 (Simulated), 81.2 (Actual)

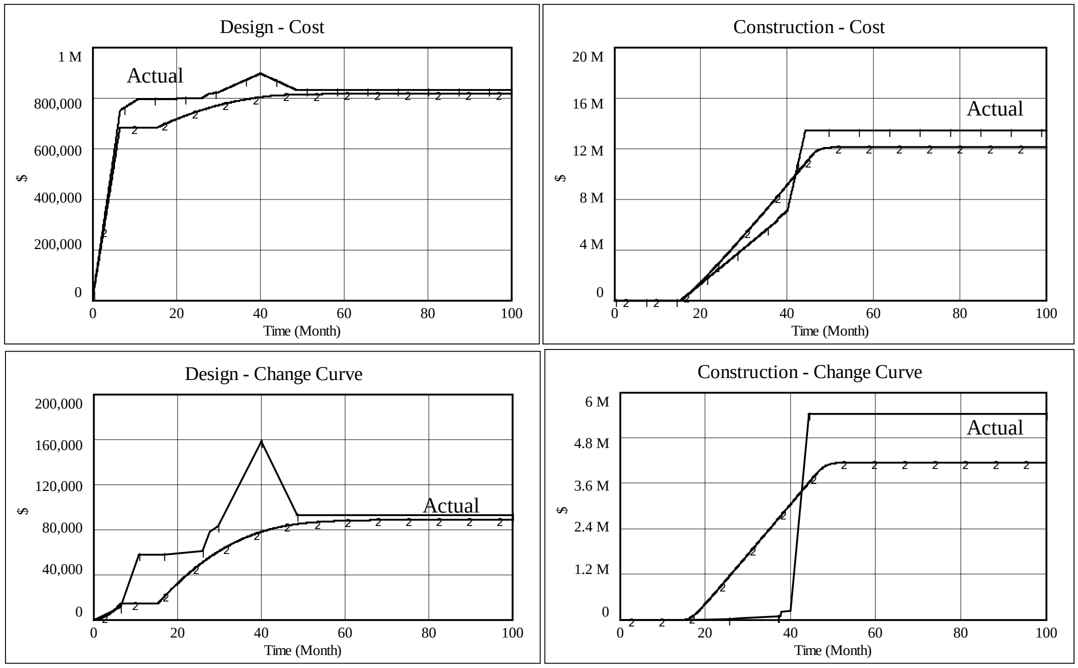

Figure 7: Simulation result of Project P040 (Worst fit). Design CD Finish = 6.5 (Simulated), 6.6 (Actual). Construction Finish = 48.7 (Simulated), 44.3 (Actual)

Robustness of Calibration Results

We conduct two sets of sensitivity analysis to assess the robustness of calibration results. First we evaluate the confidence we can have in the values reported for the feedback parameters a, b and c. Specifically, we change these parameters around their estimated value and measure the fractional change in the payoff. In the absence of formal maximum likelihood interpretation for the payoff function, we heuristically use a 20% change in payoff as a threshold that signals incongruence between the parameters and the data[4]. The results (See figure 8) suggest rather tight ranges of ±30%, ±20% and ±10% for

[4] While the complex non-parametric structure of the distributions rule out theoretical proofs, we think the 20% threshold is conservative. For demonstration, consider a maximum likelihood based payoff function with normally distributed errors (which, similar to our setting, leads to normalized squared error terms in the log-likelihood function). For a sample with N effective data points (e.g. total data points minus the number of parameters), the range of a typical log-likelihood function at the best fit position is (roughly speaking, being a chi-square distribution with N degrees of freedom) around N. In such setting, depending on the confidence levels required, a reduction of approximately 4 units in the log-likelihood (i.e. 4/N in fractional terms) signifies reasonable confidence intervals. With an N value well above 100 in our setting

these three theoretically important parameters: a, b and c respectively. The sensitivity analysis also suggests the most robust parameter is c (i.e. effect of construction progress on rework discovery rate) followed by b (effect of design quality on construction productivity).

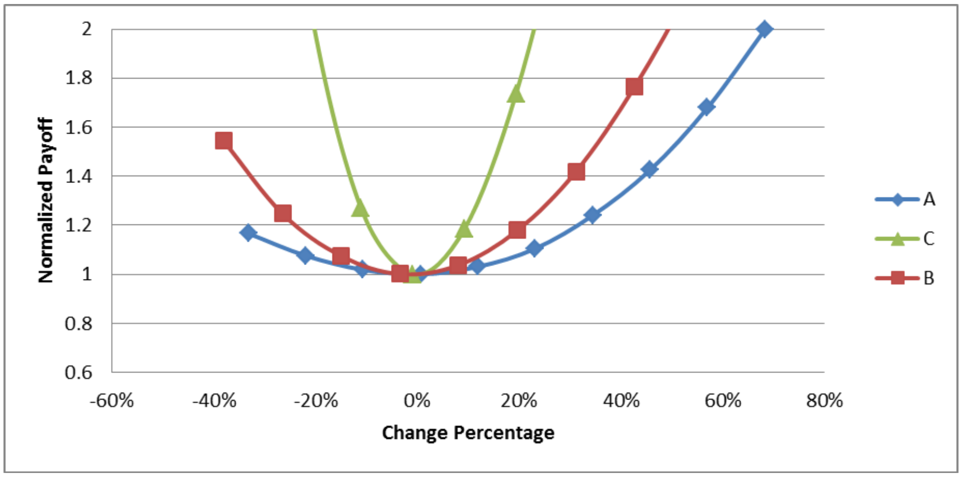

Figure 8: Significance of Parameters a, b, and c in payoff function

A second sensitivity analysis is conducted to assess the weighting functions used in defining the calibration payoff. Ten (10) different scenarios are defined with different set of weights listed in Table 3. Model is calibrated again under each scenario. The impact of error weight on different scenarios is calculated by the average of absolute percentage change of calibrated parameters. The result shows only 1% to 7% variation on calibrated parameters as the result of different error weight scenarios across different scenarios. These findings suggest that the calibration results are reasonably robust to the payoff weights used.

(15 projects multiplied by 5 single data points and 2 time series of approximately 10 data points each for each project, minus the 93 parameters estimated) we could feel confident that a fractional change in payoff of 20% (5 times the upper bound on 4/N) is fairly conservative and thus we can feel comfortable that the true values for these parameters fall within the given range.

Table 3: Scenarios of error weight sensitivity analysis

Scenario	W_{D_CD}	W_{D_DS}	W_{D_CT}	W_{D_CC}	W_{C_F}	W_{C_CT}	W_{C_CC}
1	1	1	1	1	1	1	1
2	1	1	2	1	1	2	1
3	1	1	1	2	1	1	2
4	20	20	1	1	2	1	1
5	2	2	10	1	2	1	1
6	2	2	1	10	2	1	1
7	2	2	1	1	20	1	1
8	2	2	1	1	2	10	1
9	2	2	1	1	2	1	10
10	20	20	10	10	2	1	1

Inter-phase Project Feedback Effects

Figure 9 and Figure 10 show the impacts of Factors a, b and c on construction error rate (C_E), construction production rate (C_P) and time to detect undiscovered design rework (D_D) using the calibrated values (a=2.169, b=2.232 and c=1.104). Looking at the simulations pertaining to calibration and validation project data reveals that the fraction of undiscovered design error on initial design work does not exceed the range of ±20%, which confines impact factors A and B up to 50%. Factor C's input, construction progress, ranges on the full scale of 0 to 1 which can half the rework discover time in design phase.

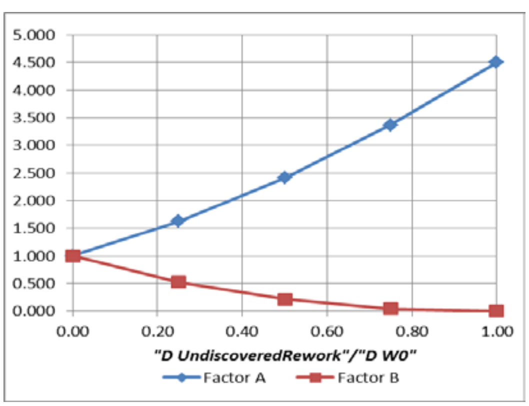

Figure 9: Factor A and B

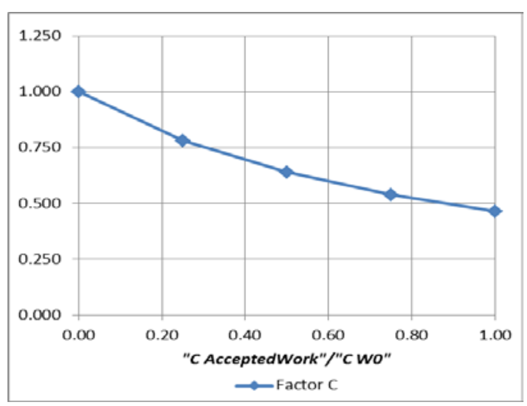

Figure 10: Factor C

Predicting the Performance of New Projects

The parameters estimated in the previous section can be used to forecast the trajectory of a completely new project before it has started. A simplistic approach is to use the average of the 15 set of calibrated parameters for the parameters of the prediction model. This approach, however, ignores the significant variability observed in parameters across different projects. Ignoring the variability would give more confidence to the projections than is warranted and deprives the user from the much valuable information regarding the expected distribution of potential performance outcomes. To address these concerns, we use a more realistic approach which assumes the six project-specific parameters are random variables with a given mean and covariance structure, available from our estimated parameters. We will then generate 1000 samples with the same mean and covariance matrices for these six (6) parameters, 3 project-specific parameters for 2 sets of design and construction, using the variance-covariance method. We assume the parameters are correlated and normally distributed. The set of random values, R, is produced by uniform random values between 0 and 1, $R_{0\text{-}1}$, using Equation 10. Matrix U is the square root of covariance matrix, $\sum$[5], calculated by Cholesky decomposition method. Table 2 reports the mean, standard deviation, and correlation matrix used for this analysis.

$$[R]=[\mu]+[U]*[R_{0\text{-}1}] \tag{10}$$

[5] $[\sum] = [D][\rho][D]$ where; $[\rho]$=correlation matrix, $[D]$ = Diagonal(σ), and σ=Standard deviation vector.

Where: $[\sum]=[U]^T\ [U]$ (11)

Next, a Monte-Carlo simulation generates the distribution of model outcomes using sample R and with a given plan (i.e. D_C0, C_C0, D_T0, and C_T0). Figure 11 to Figure 16 show the simulation result for an example project with the initial scope of $1.4M (D_C0) and $17.7M (C_C0), and scheduled duration of 12.8 months (D_T0) and 17.4 months (C_T0), for design and construction respectively. Initial scope and schedule are typically available at the beginning of any project, but are unreliable and often underestimate the actual costs and schedule significantly. These estimates are the only project-specific inputs we need in our model to generate the predicted performance projections for a new project. The project above is simulated with the 1000 sets of randomly generated parameters discussed above. 15% of the samples were found infeasible as they did not result in design and construction completion in a reasonable amount of time.

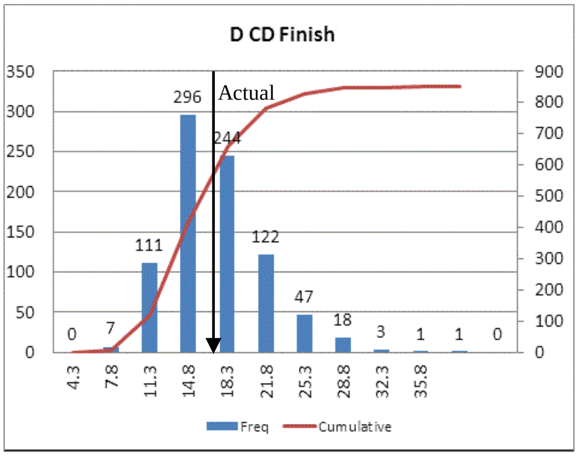

Figure 11: Distribution of design finish, Mean=15.5, StDev=4.20, Actual=18.1(month)

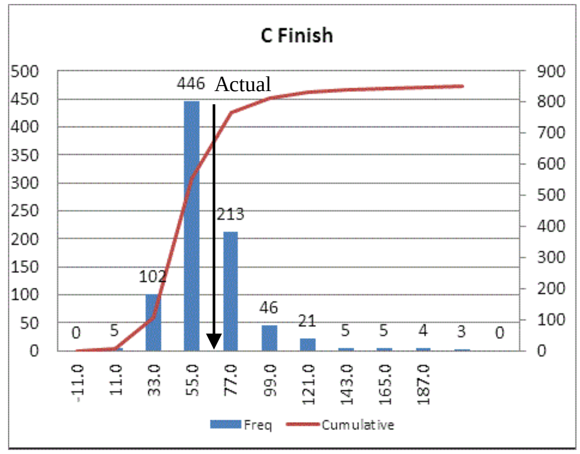

Figure 12: Distribution of construction finish, Mean=53.0, StDev=23.93, Actual=65.9(month)

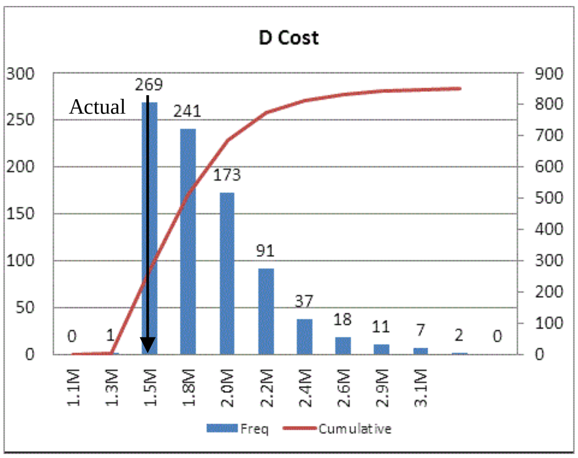

Figure 13: Distribution of design final cost, Mean=$1.7M, StDev=$0.33M, Actual=$1.5

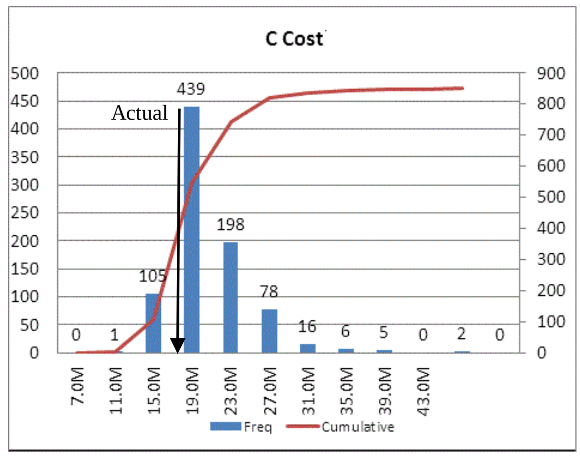

Figure 14: Distribution of construction final cost, Mean=$18.6M, StDev=$4.04M, Actual=$18.5M

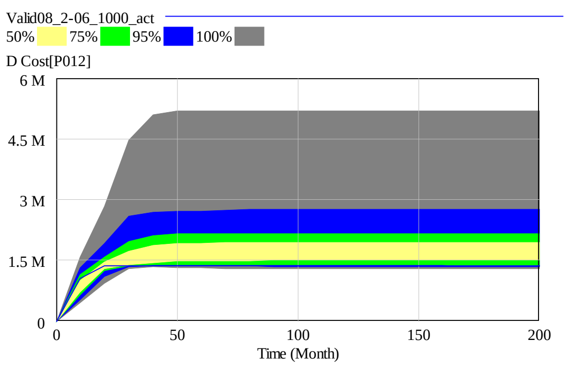

Figure 15: Distribution of design cost curve

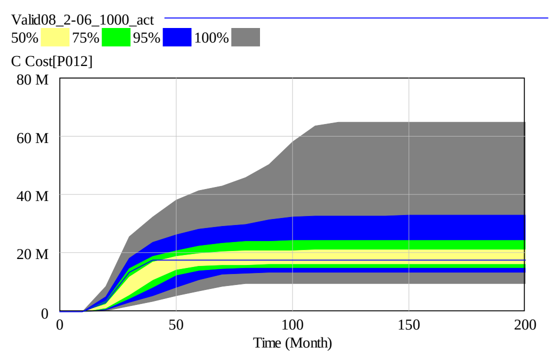

Figure 16: Distribution of construction cost curve

Validation: How well new projects can be predicted?

Building on the idea above, we can now more formally assess the ability of the model to predict the actual performance of new projects, given their original scope and schedule. Specifically, we repeat the Monte-Carlo process above for the 15 validation projects, using the 850 feasible random parameter sets. We consider four metrics including construction document finish time (D_F), design cost (D_C), construction finish time (C_F), and construction final cost (C_C). The distribution of samples produced by the Monte-Carlo simulation should be compared with the actual values for each metric and each project. For ease of comparison in each project, these simulated metrics are normalized against the actual values in that project so that the value one represents the true value. These results are reported in Figure 17 to Figure 20 in boxplot format. The box represents interquartile range (IQR) which is the distance from first (25^{th} percentiles) and third quartiles (75^{th} percentiles). The whiskers identify the maximum and minimum values. The plus symbol in the box interior represents the mean and the horizontal line in the box interior represents the median. The solid line at value 1 represents the true value.

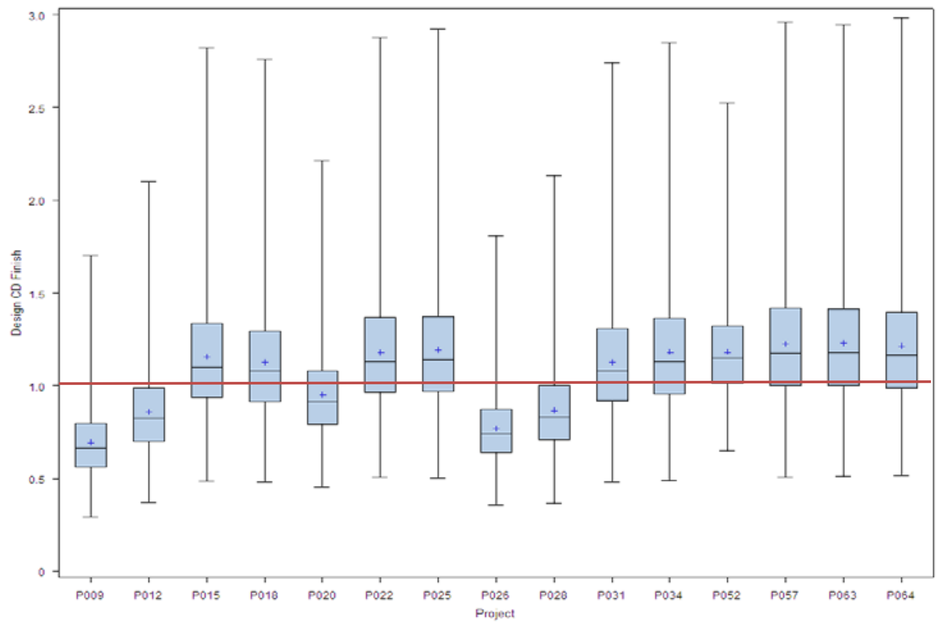

Figure 17: Boxplot of normalized design construction document finish (D_F)

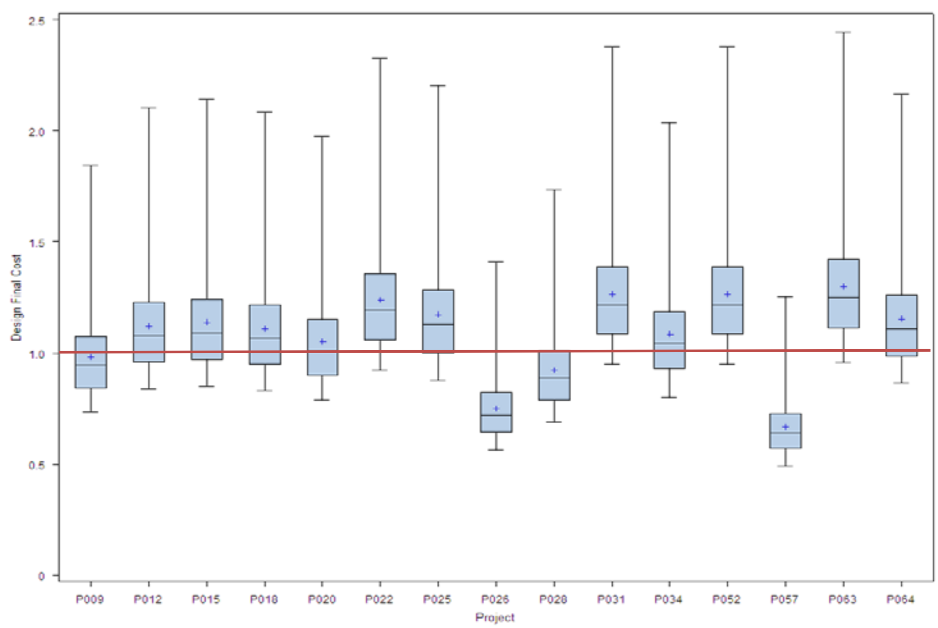

Figure 18: Boxplot of normalized design final cost (D_C)

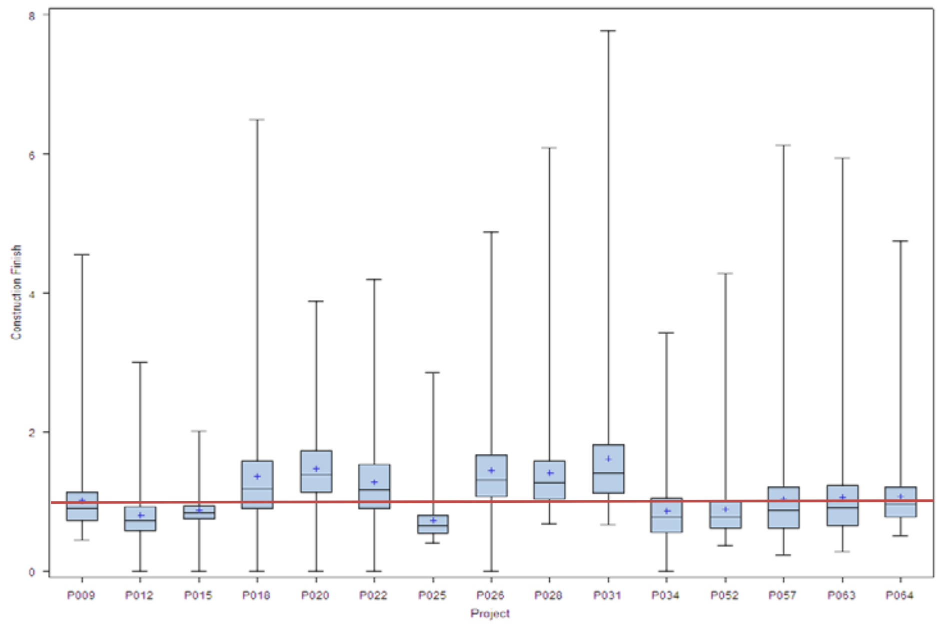

Figure 19: Boxplot of normalized construction finish (C_F)

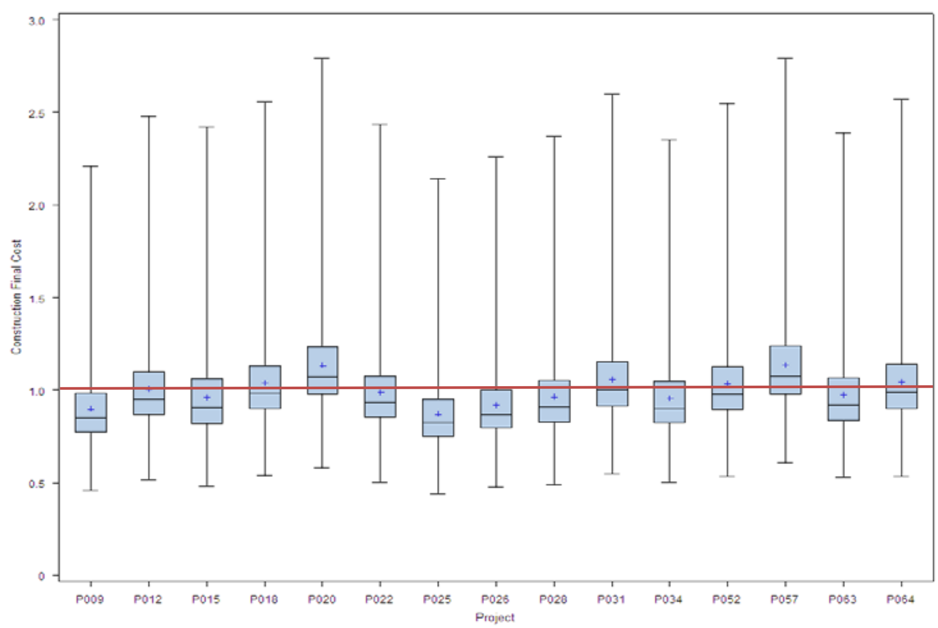

Figure 20: Boxplot of normalized construction final cost (C_F)

The best predictive model is the one which not only gets the performance measures correctly in average (i.e. no bias in the mean across many samples), but also correctly estimates the variability expected in performance. For example the model would have been over-estimating the variance if the model's mean performance always matched the actual numbers (i.e. all boxes were set squarely on value one), because the projected variability in outcomes was not borne out by the data. On the other hand, if most boxes were above, or below, the line one, we would identify a bias in the model's predictions. To better assess the overall fit of the projected model metrics against the validation data, we create a variant of Q-Q plot which combines the data from all four metrics and 15 projects into a single diagnostic graph. Consider n=60 (=15*4) actual metrics and their corresponding simulated distributions obtained through the Monte-Carlo results above. First we find what percentile each data point belongs to on the corresponding simulated distribution. The resulting data set includes 60 data points with different percentile values. We sort this dataset in the ascending order of percentiles and graph its values on x-axis against the y-axis of k/(n+1) for data point k (See Figure 21). A perfect match will be on the 45 degree line, where the empirical metrics match the corresponding percentiles in simulation distributions exactly. A bias is identified if the graph is generally above or under the 45 degree line. A line steeper than 45 degree suggests the model is over-estimating the variation in the actual metrics, i.e. it proposes many far-fetched values are possible, which actually never materialize in practice. Conversely, a less steep line than 45 degree signals the model's overconfidence in projecting as unlikely the values that are seen regularly in practice.

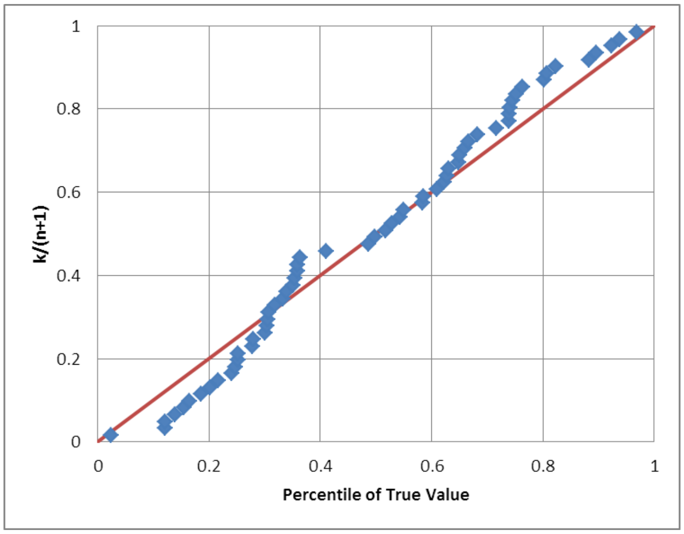

Figure 21: Q-Q plot, Percentile of true value against uniform distribution

The linear regression analysis performed on the Q-Q data series shows a very good fit between the data and regressed line with R^2 of 0.99. Moreover, 0.06 and 0.14 units of deviations are found compared to the perfect theoretical values of intercept (0) and slope (1), respectively. While no bias is found for parameter estimates (i.e. at 50th percentile), these deviations are statistically significant, suggesting a steeper slope (i.e. model slightly over predicting the variance in the outcomes). This analysis shows that while the model is pretty close in predicting the distribution of empirical final metrics, it does predict a slightly fatter tail for these metrics, than empirically observed. This suggests the model predictions are slightly pessimistic in over-predicting variation.

Table 4: Regression analysis result of Q-Q data series

	Coefficients	*Standard Error*	*P-value*
Intercept	-0.064	0.010	0.000
Slope	1.143	0.018	0.000

Conclusions

This work is the continuation of our previous work (Parvan, Rahmandad, & Haghani, 2012). In this paper, we extended the design-construction feedback relationships to three: 1) Undiscovered design rework may increase construction error rate, 2) Undiscovered design rework may slow down construction production rate, and 3) Construction progress increase the detection rate of undiscovered design reworks. We measured these feedback relationships using the empirical data of 30 construction projects. The proposed system dynamics model is in line with the previous research (Ford & Sterman, 1998; Richardson G. P. and Pugh, 1981) and provides empirical estimates for some of the important feedback mechanisms discussed in the literature. These empirical estimates validate much qualitative hypotheses in this domain and suggest the inter-phase feedback mechanisms on quality, productivity, and rework discovery time are important and at a magnitude that can make a significant impact on project dynamics.

Besides the estimation work, through our validation tests, the model was found to be a promising tool to simulate and predict construction projects. The model performed very well to match calibration sample projects. The performance of the calibrated model to predict validation sample project was fine, though it slightly over-estimated variation of outcomes. The small sample size in calibration may have led to unreliable random samples used for Monte-Carlo simulations, which would have then over-estimated the outcome variability.

In our modeling and estimation we did not consider many potentially relevant factors such as project size, project type (new/renovation), location and project complexity that may impact the project behavior and cost curve, and moderate the feedback effects of interest in our setting. Predictions may become more precise, if such data was available and used in the calibration-validation process.

Acknowledgement

We would like to sincerely thank Mr. William Clark, assistant director, Mr. Robert Martinazzi, assistant director, and Mr. Enrique Salvador, associate director, in the department of capital projects at University of Maryland for dedicating their precious time to provide comments and feedbacks which were crucial to the improvement of this research.

References:

Abdel-Hamid TK. (1984). *The dynamics of software development project management: an integrative system dynamics perspective*. MIT Sloan School of Management, Cambridge, MA.

Burati, J. L., Farrington, J. J., & Ledbetter, W. B. (1992). Causes of Quality Deviations in Design and Construction. *Journal of Construction Engineering and Management*, *118*(1), 34–49. doi:10.1061/(ASCE)0733-9364(1992)118:1(34)

Cusack, D. (1992). Implementation of ISO 9000 in Construction. *Proc., ISO 9000 Forum Symp., Institution of Engineers, Gold Coast, Australia.*

Ford, D. N., Anderson, S. D., Damron, A. J., De Las Casas, R., Gokmen, N., & Kuennen, S. T. (2004). Managing Constructibility Reviews to Reduce Highway Project Durations. *Journal of Construction Engineering and Management*, *130*(1), 33–42. doi:10.1061/(ASCE)0733-9364(2004)130:1(33)

Ford, D. N., & Sterman, J. D. (1998). Dynamic modeling of product development processes. *System Dynamics Review*, *14*(1), 31–68. doi:10.1002/(SICI)1099-1727(199821)14:1<31::AID-SDR141>3.0.CO;2-5

Hanna, A. S., Asce, M., & Gunduz, M. (2004). Impact of Change Orders on Small Labor-Intensive Projects, (October), 726–733. doi:10.1061/(ASCE)0733-9364(2004)130

Hanna, A. S., Camlic, R., Peterson, P. A., & Nordheim, E. V. (2002). Quantitative Definition of Projects Impacted by Change Orders. *Journal of Construction Engineering and Management*, *128*(1), 57. doi:10.1061/(ASCE)0733-9364(2002)128:1(57)

Hanna, A. S., Russell, J. S., Gotzion, T. W., & Nordheim, E. V. (1999). Impact of Change Orders on Labor Efficiency for Mechanical Construction. *Journal of Construction Engineering and Management*, *125*(3), 176–184. Retrieved from http://cedb.asce.org/cgi/WWWdisplay.cgi?117244

Ibbs, B. C. W. (1997). QUANTITATIVE IMPACTS OF PROJECT CHANGE : SIZE ISSUES, (September), 308–311.

Ibbs, W. (2005). Impact of Change's Timing on Labor Productivity. *Journal of Construction Engineering and Management*, *131*(11), 1219–1223. Retrieved from http://cedb.asce.org/cgi/WWWdisplay.cgi?149266

Khoueiry, Y., Srour, I., & Yassine, A. (2013). An optimization-based model for maximizing the benefits of fast-track construction activities. *Journal of the Operational Research Society*. doi:10.1057/jors.2013.30

Lee, S., Han, S., & Peña-Mora, F. (2009). Integrating construction operation and context in large-scale construction using hybrid computer simulation. *Journal of Computing in Civil …*, (April), 75–83. Retrieved from http://ascelibrary.org/doi/abs/10.1061/(ASCE)0887-3801(2009)23:2(75)

Lopez, R., & Love, P. E. D. (2012). Design Error Costs in Construction Projects, (May), 585–593. doi:10.1061/(ASCE)CO.1943-7862.0000454.

Lyneis, J. M., Cooper, K. G., & Els, S. A. (2001). Strategic management of complex projects: a case study using system dynamics. *System Dynamics Review*, *17*(3), 237–260. doi:10.1002/sdr.213

Lyneis, J. M., & Ford, D. N. (2007). System dynamics applied to project management : a survey , assessment , and directions for future research, *23*(2), 157–189. doi:10.1002/sdr

Moselhi, O., Assem, I., & El-Rayes, K. (2005). Change Orders Impact on Labor Productivity. *Journal of Construction Engineering and Management*, *131*(3), 354–359. doi:10.1061/(ASCE)0733-9364(2005)131:3(354)

Park, M., Kim, W.-Y., Lee, H.-S., & Han, S. (2011). Supply chain management model for ready mixed concrete. *Automation in Construction*, *20*(1), 44–55. doi:10.1016/j.autcon.2010.07.005

Parvan, K., Rahmandad, H., & Haghani, A. (2012). Estimating the impact factor of undiscovered design errors on construction quality. *In the proceedings of the 30th International Conference of the System Dynamics Society, St. Gallen, Switzerland*, (c), 1–16. Retrieved from http://www.systemdynamics.org/conferences/2012/proceed/papers/P1298.pdf

Rahmandad, H., & Hu, K. (2010). Modeling the rework cycle: capturing multiple defects per task. *System Dynamics Review*, *26*(4), 291–315. doi:10.1002/sdr.435

Rahmandad, H., & Sterman, J. D. (2012). Reporting guidelines for simulation-based research in social sciences. *System Dynamics Review*, *28*(4), 396–411. doi:10.1002/sdr.1481

Richardson G. P. and Pugh, A. L. I. (1981). *Introduction to system dynamics modeling with dynamo*. Cambridge, MA: Productivity Press.